On the Penrose Inequality for general horizons

Edward Malec
Instytut Fizyki, Uniwersytet Jagielloński, Reymonta 4, P- 30-059 Kraków, Poland

Marc Mars
Max-Planck-Institut für Gravitationsphysik, Am Mühlenberg 1, D-14 476 Golm, Germany

Walter Simon
Institut für theoretische Physik der Universität Wien, Boltzmanngasse 5, A-1090 Wien, Austria

For asymptotically flat initial data of Einstein's equations satisfying an energy condition, we show that the Penrose inequality holds between the ADM mass and the area of an outermost apparent horizon, if the data are restricted suitably. We prove this by generalizing Geroch's proof of monotonicity of the Hawking mass under a smooth inverse mean curvature flow, for data with non-negative Ricci scalar. Unlike Geroch we need not confine ourselves to minimal surfaces as horizons. Modulo smoothness issues we also show that our restrictions on the data can locally be fulfilled by a suitable choice of the initial surface in a given spacetime.

04.20.-q 04.30.Nk 04.40.-b 95.30.Sf

An important issue in General Relativity is the "cosmic censorship conjecture" which reads, roughly speaking, that singularities (which necessarily develop in gravitational collapse in the future of apparent horizons) are always separated from the outside world by event horizons. Penrose gave a heuristic argument which showed that, for collapsing shells of particles with zero rest mass, cosmic censorship would imply the inequality

$$M \geq \sqrt{\frac{A}{16\pi}} \tag{1}$$

between the total mass (the ADM-mass) M of the spacetime and the area A of an apparent horizon $\mathcal{H}$ [1]. While for such shells inequality (1) was recently proven by Gibbons [2], there has remained the challenge of proving (independently of cosmic censorship) the Penrose inequality (PI) (1) in the "general case" characterized below.

We consider a spacetime $(\mathcal{M}, {}^4g_{\mu\nu})$ and a corresponding initial data set $(\mathcal{N}, g_{ij}, k_{ij})$, i.e. a smooth 3-manifold $\mathcal{N}$ (which can be embedded in $\mathcal{M}$), a positive definite metric g_{ij} and a symmetric tensor k_{ij} (the second fundamental form of the embedding). To be compatible with Einstein's equations on $\mathcal{M}$, these quantities satisfy the constraints

$$D_i \left(k^i_l - k\delta^i_l\right) = -8\pi j_l, \tag{2}$$

$$R - k_{ij}k^{ij} + k^2 = 16\pi\rho. \tag{3}$$

where D_i is the covariant derivative and R the Ricci scalar on $\mathcal{N}$, $k = g^{ij}k_{ij}$, and ρ and j_i are the energy density and the matter current, respectively. We take the data to be asymptotically flat and to satisfy the dominant energy condition (which implies that $\rho \geq |j|$). Furthermore, we assume that the boundary of $\mathcal{N}$ (if non-empty) is a "future apparent horizon" $\mathcal{H}$ (called a "horizon" from now on) which is a 2-surface defined by the property that all outgoing future directed null geodesics (in $\mathcal{M}$) orthogonal to $\mathcal{H}$ have vanishing divergence, i.e. $\theta_+ = 0$, and that the divergence of the outgoing past null geodesics is non-negative, i.e. $\theta_- \geq 0$. (Past apparent horizons are defined similarly, and satisfy analogous theorems). If $\theta_+ = \theta_- = 0$ on $\mathcal{H}$, the horizon is an extremal surface of $(\mathcal{N}, g_{ij})$ and the outermost extremal surface in an asymptotically flat space must be minimal.

As an idea for proving the positive mass theorem (PMT) Geroch considered an asymptotically flat Riemannian manifold $(\mathcal{N}, g_{ij})$ with non-negative Ricci scalar R (which naturally arises from initial data sets described above by restricting k_{ij} suitably) and assumed that on $(\mathcal{N}, g_{ij})$ there is a smooth "inverse mean curvature flow" (IMCF). This means that one can write the metric g_{ij} as

$$ds^2 = \phi^2 dr^2 + q_{AB}dx^A dx^B \tag{4}$$

(where $A, B = 2, 3$), with smooth fields ϕ and q_{AB}, and with $p\phi = 1$ where p is the mean curvature of the surfaces $r = \text{const.}$ (This parametrization yields $dA/dr = A$, i.e. the area A of the surfaces increases exponentially). Assuming also that these surfaces have spherical topology, Geroch showed that the mass functional

$$M_G(\mathcal{S}) = \frac{\sqrt{A}}{64\pi^{\frac{3}{2}}}\left(16\pi - \int_{\mathcal{S}} p^2 dS\right) \tag{5}$$

is monotonic under a smooth IMCF, i.e. $dM_G/dr \geq 0$ [3]. In some cases (in particular, when $k_{ij} = 0$) M_G is a special case of the functional

$$M_H(\mathcal{S}) = \frac{\sqrt{A}}{64\pi^{\frac{3}{2}}}\left(16\pi - \int_{\mathcal{S}} \theta_+\theta_- dS\right) \tag{6}$$

introduced earlier by Hawking [4]. Recently, Huisken and Ilmanen proved that monotonicity of M_G also holds without the smoothness assumption on the flow [5,6]. Since M_G tends to M at spatial infinity I^0, flowing by IMC out of a point proves the positive mass theorem ($M \geq 0$) on

manifolds without minimal surfaces (for which $p = 0$). On the other hand, since M_G equals $\sqrt{A/16\pi}$ on a minimal surface, flowing from the outermost such surface to infinity proves the special case of the PI (1) for which horizons coincide with minimal surfaces [5]- [7].

In the present Letter we show that the Hawking mass (6) is also monotonic under a smooth IMCF provided that certain rather simple supplementary conditions are satisfied. To motivate the latter and to compare them with previous work, it is useful to formulate first, as a conjecture, the sharpened version

$$M^2 \geq \frac{A}{16\pi} + P^2. \tag{7}$$

of the PI (c.f. [13] for a careful formulation). This inequality seems natural in view of the PMT $M \geq |P|$ for the ADM-momentum P [9,10] (which also holds in the presence of apparent horizons [11,12]). Note that the anticipated result (7) only involves the area A and the norm $M^2 - P^2$ of the ADM 4-vector P^μ which are *spacetime quantities*. On the other hand, we wish to show monotonicity of (6) for a class of *data* $(\mathcal{N}, g_{ij}, k_{ij})$ as general as possible. For this purpose we can now pursue either an "invariant" approach in which an arbitrary hypersurface $\mathcal{N}$ connecting $\mathcal{H}$ with I^0 is admitted, or a "gauge" approach for which $\mathcal{N}$ is fixed suitably. While in the former setting we obtain monotonicity of M_H only for a restricted class of data, this monotonicity then directly implies the PI since (like M_G) M_H still tends to $\sqrt{A/16\pi}$ on the horizon and to M at I^0. On the other hand, in the "gauge approach" we can implement restrictions on the data by determining $\mathcal{N}$ via initial conditions on $\mathcal{H}$ and via a suitable *local* propagation law. However, here we cannot guarantee that $\mathcal{N}$ really ends up at I^0 rather than becoming hyperboloidal and approaching null infinity.

As to the "invariant" approach, Jang has presented a generalization of the functional (5) which is, for *any* data $(\mathcal{N}, g_{ij}, k_{ij})$, monotonic when propagated outwards with a flow determined by a certain quasilinear partial differential equation [14]. If existence of solutions for this equation could be proven, then it would follow that $M \geq |P|$, and $M = 0$ would imply that $(\mathcal{N}, g_{ij}, k_{ij})$ are data for flat space. Unfortunately, Jang's equation is too complicated to be tractable at present. Moreover, as to possible extensions towards a PI, the boundary terms at the horizon seem to be difficult to analyze in this method.

On the other hand, the "gauge" approach to proving the PI has been put forward by Frauendiener [8]. He suggests to construct a foliation $\mathcal{S}(r)$ and a hypersurface $\mathcal{N}$ by starting from $\mathcal{H}$ and moving outwards with the "inverse mean curvature vector" (assuming this is spacelike),

$$J^\mu = \frac{1}{\theta_+} l_+^\mu + \frac{1}{\theta_-} l_-^\mu = \frac{\sqrt{2}}{\theta_+ \theta_-} \left(p m^\mu - q n^\mu \right), \tag{8}$$

where l_+^μ and l_-^μ are, respectively, tangent to the future and past outgoing null geodesics emanating orthogonally from a level surface $\mathcal{S}(r)$, with $l_+^\mu l_{-\mu} = 1$, $\sqrt{2} m^\mu = l_+^\mu + l_-^\mu$, $\sqrt{2} n^\mu = l_+^\mu - l_-^\mu$, $2p = \theta_+ + \theta_-$ and $2q = \theta_+ - \theta_-$. While J^μ is uniquely defined (by the first equation (8)), m^μ, n^μ, p and q depend on the scaling $l_+^\mu \to \lambda l_+^\mu$, $l_-^\mu \to \lambda^{-1} l_-^\mu$ (with some function λ on $\mathcal{S}(r)$). By choosing λ suitably we can achieve that m^μ and n^μ are, respectively, parallel and orthogonal to any spacelike hypersurface $\mathcal{N}$. In particular, when J^μ is required to be tangent to $\mathcal{N}$, we have $q \equiv 0$ there. This so-called "polar hypersurface condition" [15] implies $R \geq 0$, and it also implies that $\mathcal{N}$ can reach only those horizons (where $\theta_+ = 0$) which also satisfy $p = 0$ and are thus minimal surfaces. Therefore Frauendiener's approach in essence boils down to the situation considered by Geroch already [3]. An advantage of Frauendiener's 4-dimensional reformulation is that he can set out from the "weak energy condition" to *obtain* $R \geq 0$ instead of *imposing* the latter condition. However, as mentioned before, it is not clear whether the constructed hypersurface really reaches I^0.

In the theorem below, we prove the "weak form" (1) of the PI for "general" future apparent horizons, within the "invariant aproach" and for a restricted class of data. We will then show that, modulo smoothness issues, these restricions can be omitted within a "gauge" approach where, in fact, a large family of hypersurfaces is admitted.

The following definitions apply to a smooth (but otherwise arbitrary) foliation of $\mathcal{N}$ by 2-surfaces $\mathcal{S}(r)$. Let m^i denote the unit normal to $\mathcal{S}(r)$, ϕ the "lapse" (c.f. (4)), $q_{ij} = g_{ij} - m_i m_j$ the induced metric, p_{ij} the second fundamental form of $\mathcal{S}$ in $\mathcal{N}$, t_{ij} its trace-free part (i.e. $t_{ij} = p_{ij} - (p/2) q_{ij}$) and p its trace (the mean curvature). We will also employ the following decomposition of k_{ij} w.r. to $\mathcal{S}(r)$,

$$k_{ij} = z m_i m_j + m_i s_j + m_j s_i + q_i^k q_j^l x_{kl} + \frac{1}{2} q q_{ij}, \tag{9}$$

where $z = k_{ij} m^i m^j$, $s_i = q_i^j k_{jl} m^l$, $q = k_{ij} q^{ij}$ and $x_{ij} = q_i^l q_j^n k_{ln} - (1/2) q q_{ij}$. From p and q we can define θ_+ and θ_- using the expressions following (8).

Theorem. Let $(\mathcal{N}, g_{ij}, k_{ij})$ be a smooth, asymptotically flat initial data set for Einstein's equations with an outermost future apparent horizon $\mathcal{H}$ (i.e. $\theta_+ = 0$ and $\theta_- \geq 0$ on $\mathcal{H}$) of spherical topology, which satisfies the constraints (2),(3), the dominant energy condition, and the following additional restrictions

1. On $(\mathcal{N}, g_{ij})$ there exists a smooth inverse mean curvature flow.

2. The divergence θ_- (taken with respect to the level sets of the IMCF) is positive outside $\mathcal{H}$.

3. On each level set of the IMCF, at least one of (a) or (b) holds.

(a) q/p is constant.

(b) $q^{ij}D_i s_j = 0$, and $\theta_+ \geq 0$.

Then the PI (1) holds.

Proof. We show monotonicity of the Hawking mass functional (6). Noting that $\theta_+\theta_- = p^2 - q^2$, we compute the derivatives of p and q in the direction of m^i (the IMCF is not required for this step). Using (9), the 3- and the 2-dimensional constraints and the fact that $\phi m^i D_i m_k = -q_k^j D_j \phi$ in the coordinates (4) gives

$$\begin{aligned}
2m^iD_ip &= -2\frac{{}^2\Delta\phi}{\phi} - 16\pi\rho - 2s_is^i - \\
&\qquad x_{ij}x^{ij} + \frac{1}{2}q^2 + 2zq + {}^2R - \frac{3}{2}p^2 - t_{ij}t^{ij}, \\
m^iD_iq &= 8\pi j_im^i - x_{ij}t^{ij} + p(z - \frac{1}{2}q) + \\
&\qquad q^{ij}D_is_j + 2\frac{s^iD_i\phi}{\phi},
\end{aligned} \tag{10}$$

where ${}^2R$ and ${}^2\Delta$ are the Ricci scalar and the Laplacian with respect to q_{ij}. Next we use that $m^iD_i\sqrt{g} = p\sqrt{g}$, restrict ourselves to an IMCF (i.e. $\phi p = 1$) and remove the ${}^2\Delta$-term via integration by parts. We obtain, for the Lie derivative $\mathcal{L}_{\phi m^i}$ of M_H,

$$\begin{aligned}
\mathcal{L}_{\phi m^i}M_H(\mathcal{S}) = \frac{\sqrt{A}}{64\pi^{\frac{3}{2}}}\int_{\mathcal{S}}\Bigg[&16\pi\left(\rho + \frac{q}{p}j_im^i\right) + \\
&+\left(x_{ij}x^{ij} - 2\frac{q}{p}x_{ij}t^{ij} + t_{ij}t^{ij}\right) + \\
&+2\left(s_is^i - 2\frac{q}{p}s^i\frac{D_ip}{p} + \frac{q^{ij}(D_ip)(D_jp)}{p^2}\right) + \\
&+2\frac{q}{p}q^{ij}D_is_j\Bigg]dS.
\end{aligned} \tag{11}$$

We first show that $|q/p| \leq 1$ which is obvious when conditions 2. and 3.(b) hold. If in turn we assume conditions 2. and 3.(a), we can write $\theta_+ = \gamma\theta_-$ where $\gamma = (1 + q/p)/(1 - q/p)$ is constant on each $\mathcal{S}$. Since $\theta_+ > 0$ at some point exterior to $\mathcal{H}$, the same is true for γ. If γ were to vanish somewhere in the exterior, it must vanish on the whole leaf $\mathcal{S}$, in which case $\mathcal{H}$ would not be the outermost horizon. Therefore, $\gamma > 0$ and $\theta_+ > 0$ in the exterior, which implies our claim.

Now the dominant energy condition together with $|q/p| \leq 1$ implies that the first term (in parentheses) on the r.h. side of (11) is non-negative. Next, again due to $|q/p| \leq 1$, the second and the third terms are positive quadratic forms. Finally, the last term vanishes obviously by condition 3.(b), by partial integration on $\mathcal{S}(r)$ due to 3.(a). Hence $\mathcal{L}_{\phi m^i}M_H \geq 0$, and integrating (11) between the horizon and infinity finishes the proof.

We remark that conditions 2. and 3. could be substituted by any weaker condition which still makes the r.h. side of (11) non-negative (in particular we could just demand this property). The reason for having selected 2. and 3. is that these conditions seem to leave enough freedom such that, in a given spacetime, they could always be satisfied by a suitable choice of the hypersurface. Moreover, they seem simple enough such that this conjecture could be proven. In fact, consider a spacetime $\mathcal{M}$ and an arbitrary (smooth) function $F(r)$ with $F(0) = -1$ and $|F(r)| < 1$ for $r > 0$. Assuming that "everything" is smooth, we can then find a hypersurface $\mathcal{N}$ such that the data induced on $\mathcal{N}$ satisfy 1., 2. and 3.(a), where $r = const$ are the level sets of the IMCF, with $\mathcal{H}$ located at $r = 0$, and with $q/p = F(r)$. To see this on a heuristic basis, assume we have found a "piece" of such a hypersurface, bounded by a level set $\mathcal{S}_0$ (given by $r = r_0$) of the IMCF. Rescaling the null vectors l_-^μ and l_+^μ emanating from $\mathcal{S}_0$ like $l'^\mu_+ = \lambda l^\mu_+$ and $l'^\mu_- = \lambda^{-1}l^\mu_-$, with $\lambda^2 = \theta_-\theta_+^{-1}(1 + F(r))(1 - F(r))^{-1}$ (and keeping λ constant along l_+^μ and l_-^μ) the spacelike vector field $\sqrt{2}m'^\mu = l'^\mu_+ + l'^\mu_-$ then determines the direction in which $\mathcal{N}$ has to be continued across $\mathcal{S}_0$. Since $F(r)$ can be chosen arbitrarily, it is also plausible that a suitable choice would make the surface $\mathcal{N}$ reach I^0.

We now turn to the issue of removing condition 1., both for "arbitrary" and for "selected" hypersurfaces. Recall that, in the "time-symmetric" context, Huisken and Ilmanen have defined a generalized ("weak") IMCF flow [5,6]. The latter "jumps" at those (countably many) parameter values for which the set swept out by the flow can be enclosed by another one of the same area but with larger volume. Due to the requirement on the area, the additional segments of the new hull necessarily consist of minimal surfaces, i.e. $p = 0$. Such segments do not contribute to M_G, whence the latter *increases* at each jump. In the more general situation considered here, these "jumps" seem still appropriate since they go well with IMCF. However, they will not directly preserve monotonicity of M_H in general. Though probably very difficult, this problem could be addressed in several ways. First, one could take a pure "initial data" perspective, in which the constraints have to be solved at the same time as the flow. In this case, the value of q after the jump is not given a priori but still needs to be obtained. Thus, it is conceivable that general enough initial data exist for which M_H is still monotonic at the jumps. Alternatively, one could take a spacetime point of view. It may be possible to generalize the procedure for defining a hypersurface sketched above for the smooth case, to obtain a weak flow such that 2. and 3. are satisfied and M_H is monotonic. Instead, it may turn out to be necessary to allow for more general flows and let the two-surfaces "jump in spacetime", by which we mean that they are not a priori restricted to *any 3-surface whatsoever.* In any case,

it seems reasonable to *conjecture* that the theorem above holds if "initial data set" is replaced by "spacetime", and if all *conditions 1. - 3. are removed*.

Some further comments are in order.

Equation (11) can be written in integral form as $M_H(\mathcal{S}_2) = M_H(\mathcal{S}_1) + M_V$. Here $M_V = \int_V B d\varrho dS/16\pi$, B stands for the integrand in (11), $\varrho = \sqrt{A/4\pi}$ is an areal radius coordinate, $\mathcal{S}_1$ and $\mathcal{S}_2$ are two leaves of the foliation and $V(\mathcal{S})$ is the volume of the annulus enclosed by $\mathcal{S}_1$ and $\mathcal{S}_2$. M_V is nonnegative under the assumptions stated in the theorem. When $\mathcal{S}_1$ reduces to a point, M_V can be regarded as a volume representation of the Hawking mass $M_H(\mathcal{S}_2)$. If the internal boundary $\mathcal{S}_1$ lies outside $\mathcal{H}$, we obtain $M \geq \sqrt{A/16\pi} + M_V$ (where A is still the area of $\mathcal{H}$). Notice that outside $\mathcal{H}$ the quantity M_V can be useful to give a bound from above to the energy norm of matter fields; this may become relevant in the investigation of the Cauchy problem.

As to known special cases of our formula (11), it reduces for $k_{ij} = 0$ to the expression obtained by Geroch [3] and for $q = 0$ (upon translating to spinor formalism) to Frauendiener's expression (equ. (9) of [8]). In the spherically symmetric case (i.e. when g_{ij} and k_{ij} are invariant under rotations) the PI was proven by Malec and O'Murchadha [16] and by Hayward [17]. This result can also be recovered from the theorem above, since the level surfaces of the IMCF are metric spheres in this case, and each of the conditions in 3. is obviously satisfied.

Compared with Jang's ideas for proving the PMT for general data, we need to restrict our data strongly, but we require only the knowledge of the IMCF rather than proving existence for an involved quasilinear PDE. We note that Jang sets out from a mass functional which is different from (6), even after imposing our gauge condition 3. Therefore, imposing the latter in Jang's computation will not yield the present results.

Our approach might also be useful to prove the PI involving the Bondi mass rather than the ADM one. Under suitable technical assumptions, this version of the PI has been obtained by Ludvigsen and Vickers [18] and by Bergqvist [19] by showing that certain mass functionals (different from (6)) are monotonic along *null hypersurfaces*, and that they have the appropriate values at the horizon and at null infinity. The same technique has been applied to the Hawking mass (6) by Hayward [20] who obtains a pair of monotonicity formulas for M_H along the two null directions. We note that by just taking linear combinations of Hayward's expressions one does *not* obtain the Lie derivative of M_H in the corresponding spacelike direction (as claimed in [20]).

Finally, we discuss the possibility of obtaining generalizations of the PI (1). For spacetimes with electromagnetic fields the inequality $M \geq |Q|$, where Q is the electric charge [11,12], has been shown (also in the presence of apparent horizons) whenever the norm of the charge 4-current is not larger than the norm of the matter 4-current. Moreover, both for the case $k_{ij} = 0$ as well as for general horizons in spherically symmetric spacetimes, it is known that $M \geq \sqrt{A/16\pi} + Q^2\sqrt{\pi/A}$ provided all charges are inside $\mathcal{H}$ [16,21]. Under the same requirements, the same generalized PI can be shown for general non-spherical horizons by applying the same arguments as in [21] to our final expression (11). One could also try to incorporate the linear momentum into the inequality. It might be possible to get (7) (or something similar) by extracting the momentum out of (11) or, perhaps better, by looking for an alternative energy functional which directly gives $M^2 - P^2$ at infinity.

Acknowledgements. E.M. was supported in part by the KBN grant 2 PO3B 010 16, and W.S. was supported by the Austrian FWF project P14621-Mat. W.S. is grateful to Helmuth Urbantke for helpful discussions.

[1] R. Penrose, Ann. N.Y. Acad. Sci. **224**, 125 (1973).
[2] G. Gibbons, Class. Quantum Grav. **14**, 2509 (1977).
[3] R. Geroch, Ann. N.Y. Acad. Sci. **225**, 108 (1973).
[4] S. Hawking, J. Math. Phys. (N.Y.) **9**, 598 (1968).
[5] G. Huisken, T. Ilmanen, Int. Math. Res. Not. **20**, 1045 (1997).
[6] G. Huisken, T. Ilmanen, J. Diff. Geom. (to be published).
[7] P.S. Jang and R.M. Wald, J. Math. Phys.(N.Y.) **18**, 41 (1977).
[8] J. Frauendiener, Phys. Rev. Lett. **87**, 101101-1 (2001).
[9] R. Schoen, S.T. Yau, Commun. Math. Phys. **79**, 231 (1981).
[10] E. Witten, Commun. Math. Phys. **80**, 381 (1981).
[11] G. Gibbons, S.W. Hawking, G.T. Horowitz, M. Perry, Commun. Math. Phys. **88**, 295 (1983).
[12] M. Herzlich, J. Geom. Phys. **26**, 97 (1998).
[13] G. T. Horowitz, in *Asymptotic Behavior of Mass and Spacetime Geometry, Spinger Lecture Notes in Physics 202*, edited by F. J. Flaherty, (Springer 1984).
[14] P.S. Jang, J. Math. Phys. **19**, 1152 (1978); erratum: J. Math. Phys. **20**, 217 (1979).
[15] J. M. Bardeen and T. Piran, Phys. Rep. **96**, 205 (1983).
[16] E. Malec, N. O' Murchadha, Phys. Rev. D **49**, 6931 (1994).
[17] S. Hayward, Phys. Rev. D **49**, 6467 (1994).
[18] M. Ludvigsen, J. A. G. Vickers, J. Phys. A: Math. Gen. **16**, 3349 (1983).
[19] G. Bergqvist, Class. Quantum Grav. **14**, 2577 (1997).
[20] S. Hayward, Class. Quantum Grav. **11**, 3037 (1994).
[21] P. S. Jang, Phys. Rev. D **20**, 834 (1979).